Construction and Exploration of the Far-Ultraviolet and Far-Infrared Bivariate Luminosity Function of Galaxies

Tsutomu T. TAKEUCHI

Division of Particle and Astrophysical Science,
Nagoya University, JAPAN

Akane SAKURAI

Division of Particle and Astrophysical Science,
Nagoya University, JAPAN

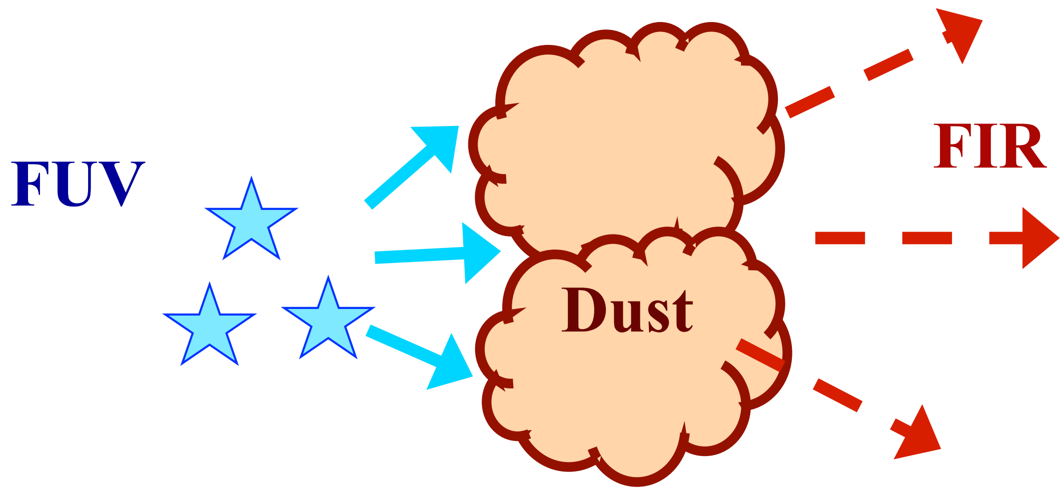
Véronique BUAT, Denis BURGARELLA

Laboratoire d'Astrophysique de Marseille, OAMP, FRANCE

From Dust to Galaxies, Paris, 27th June – 2nd July 2011

1 Evolution of the Visible and Hidden Star Formation

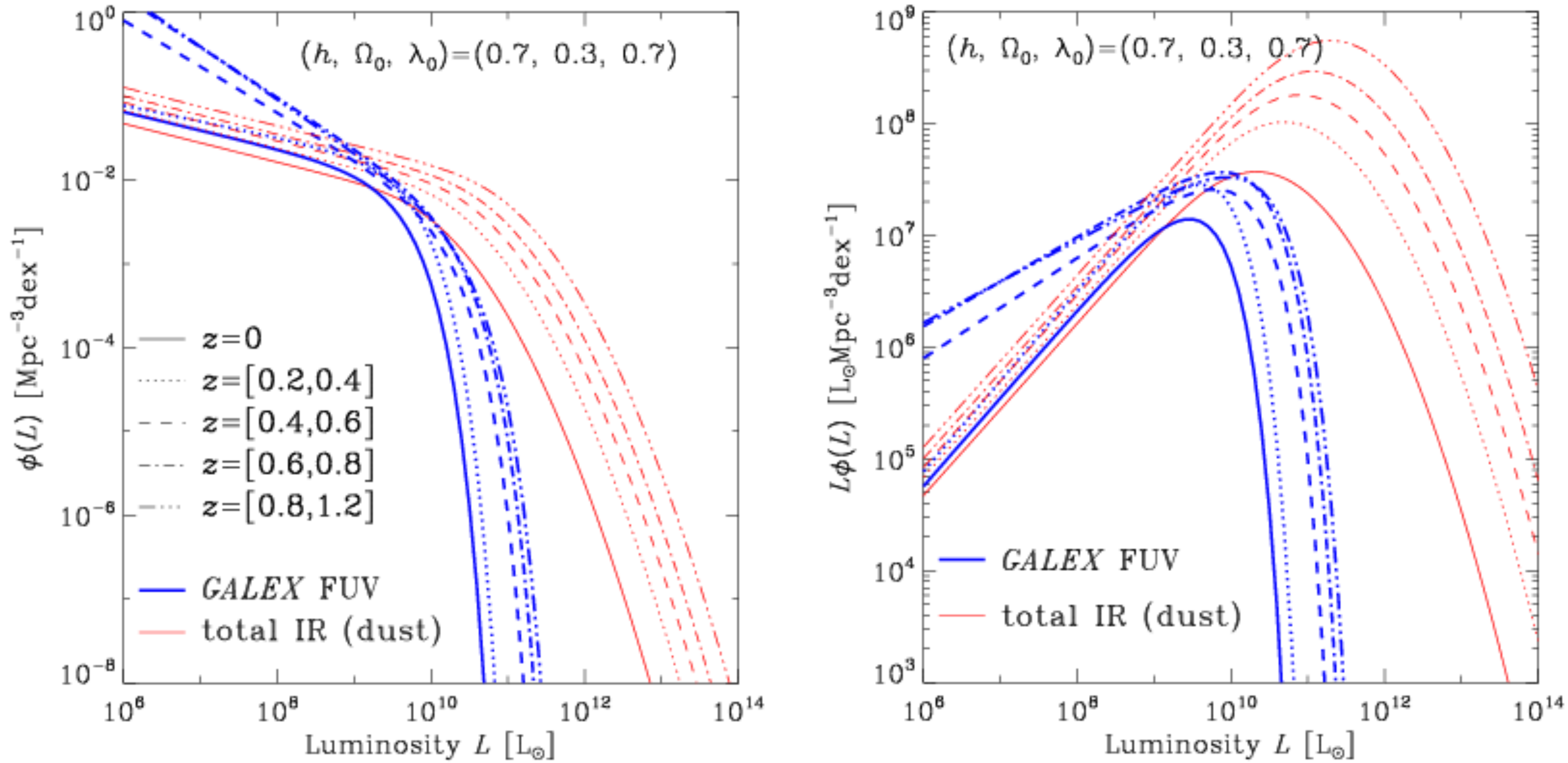
What individual analyses tell about the star formation history?



Obviously the amount of UV light absorbed by dust is only measured at FIR wavelengths. Hence, to obtain an unbiased view of the cosmic star formation, it is crucial to treat the information of both FUV and FIR (and others).

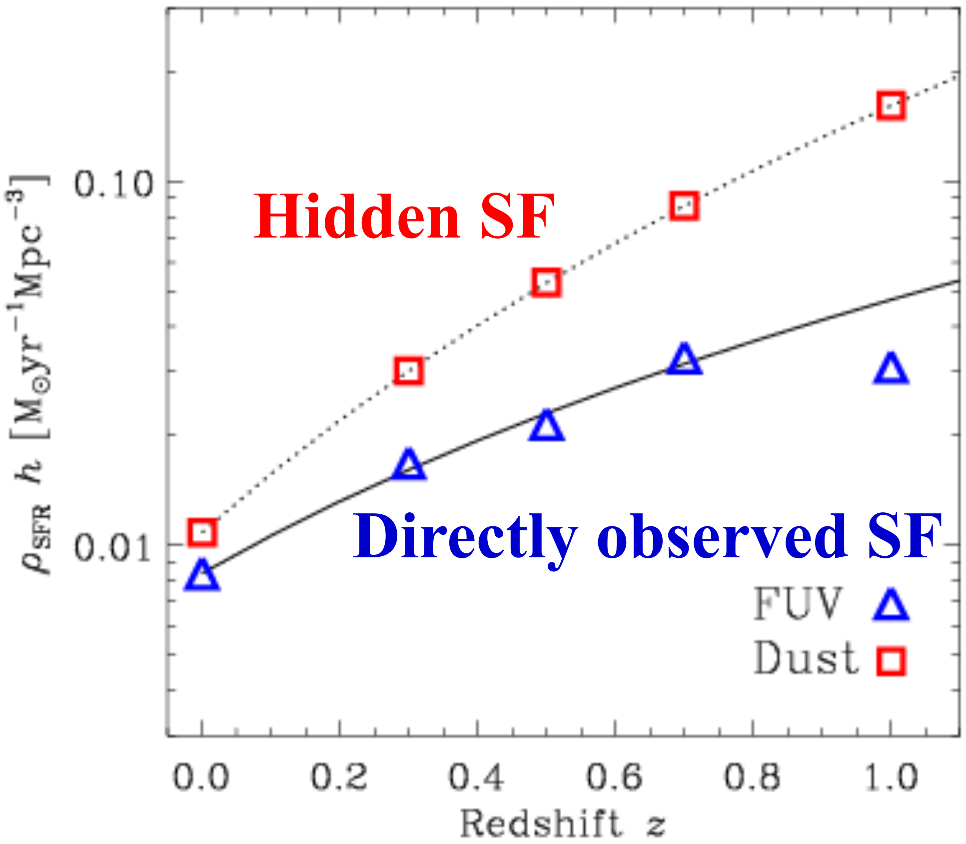
Now various multiwavelength survey data are available, and we can study the cosmic SF history a coherent and synthesized manner.

Evolution of the FUV and FIR luminosity functions



(Takeuchi, Buat, & Burgarella 2005)

Evolution of visible and hidden SF in the Universe



The local fraction of the hidden SF is 50-60%, while the fraction at z=1 reaches more than 90%.

(Takeuchi, Buat, & Burgarella 2005)

Dusty era of the Universe

Later works confirmed this "dusty era of the Universe", and revealed that the dominance of the hidden SF continues even toward higher redshifts ($z \sim 3$) (e.g., Murphy et al. 2011).

With a comparison between individual datasets from different bands, now we have a rough picture of the visible and hidden part of the cosmic SF.

Dusty era of the Universe

Later works confirmed this "dusty era of the Universe", and revealed that the dominance of the hidden SF continues even toward higher redshifts ($z \sim 3$) (e.g., Murphy et al. 2011).

With a comparison between individual datasets from different bands, now we have a rough picture of the visible and hidden part of the cosmic SF.

What is next? What does the different evolution at different wavelength mean?

To answer this question, we need to model the dependence structure between UV and IR luminosities.

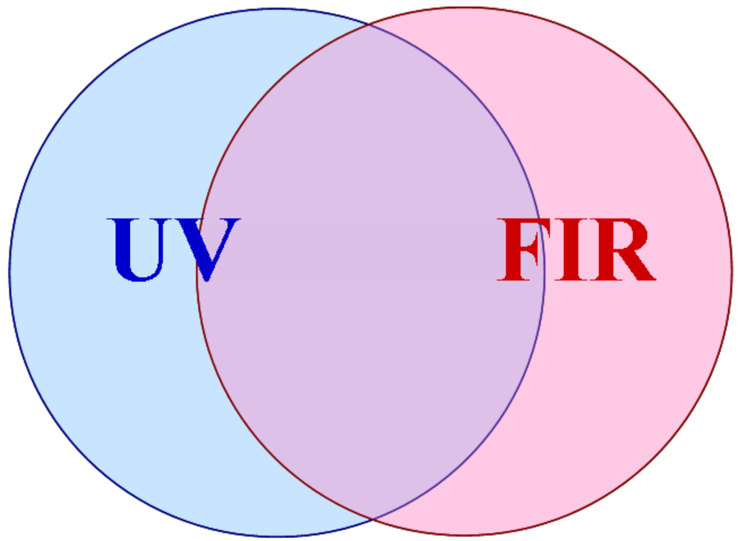
2 Bivariate Luminosity Function Analysis: Formulation

Bivariate analysis: structure of the datasets

It is very important to understand how we select sample galaxies and what we see in them. Each time we find some relation between different properties, we must understand clearly which is real (physical) and which is simply due to a selection effect.

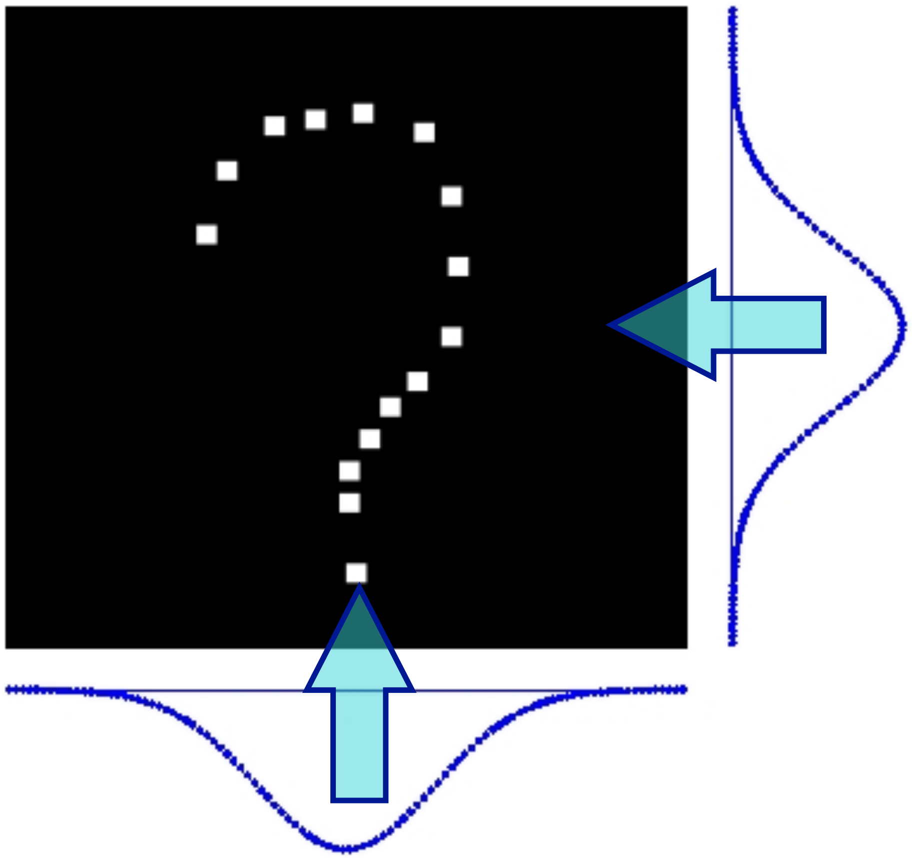
Structure of the sample is schematically described as the Venn diagram.

To handle this problem, we need a UV-IR bivariate LF.



Copula: a mathematical tool to combine marginal distributions

Question: can we (re)construct a multivariate probability density function (PDF) from its marginals?



Pougaza (2009)

Copula: a mathematical tool to combine marginal distributions

Obviously, there is an infinite number of degrees of freedom to choose the original PDF, because the dependence structure is not specified.

Then, is this problem completely unsolvable?

Copula: a mathematical tool to combine marginal distributions

Obviously, there is an infinite number of degrees of freedom to choose the original PDF, because the dependence structure is not specified.

Then, is this problem completely unsolvable?

The answer is not entirely, if we can restrict or specify the dependence between variables. The tool to deal with this problem is the copula, with a general form as follows:

$$G(x_1, x_2) = C\left[F_1(x_1), F_2(x_2)\right] \quad \textbf{(1)}$$

where $F_1(x_1)$ and $F_2(x_2)$ are two univariate marginal cumulative distribution functions (DFs) and $G(x_1, x_2)$ is a bivariate DF.

Copula: a mathematical tool to combine marginal distributions

Theorem: *Sklar's theorem*

Let G be a joint distribution function with margins F_1 and F_2. Then, there exists a copula C such that for all x_1, x_2,

$$G(x_1, x_2) = C[F_1(x_1), F_2(x_2)] \tag{2}$$

This theorem guarantees that *any bivariate DF* with given margins can be expressed with a form of equation (2). This theorem also guarantees that if we fix F_1, F_2, and the dependence structure C, the bivariate DF is uniquely determined.

Gaussian copula

Since the choice of copula is literally unlimited, we have to introduce a guidance principle.

In many data analyses in physics, the most familiar measure of dependence might be the linear correlation coefficient ρ. Mathematically speaking, ρ depends not only on the dependence of two variables but also the marginal distributions, which is not an ideal property as a dependence measure. Even so, a copula having an *explicit* dependence on ρ would be convenient.

In this work, we use a copula with this property, the Gaussian copula.

Gaussian copula

Let

$$\psi_1(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \tag{3}$$

$$\Psi_1(x) = \int_{-\infty}^{x} \psi_1(x')dx' \tag{4}$$

and

$$\psi_2(x_1, x_2; \rho) = \frac{1}{\sqrt{(2\pi)^2(1-\rho^2)}} \exp\left[-\frac{x_1^2 + x_2^2 - 2\rho x_1 x_2}{2(1-\rho^2)}\right] \tag{5}$$

$$\Psi_2(x_1, x_2; \rho) = \int_{-\infty}^{x_1}\int_{-\infty}^{x_2} \psi_1\left(x_1', x_2'\right)dx_1' dx_2' \tag{6}$$

Gaussian copula

We then define a Gaussian copula C^G (u^1, u^2; ρ) as

$$C^{\mathrm{G}}(u_1, u_2; \rho) \equiv \Psi_2\left[\Psi_1^{-1}(u_1), \Psi_1^{-1}(u_2); \rho\right] \tag{7}$$

The density of C^G, c^G, is obtained as

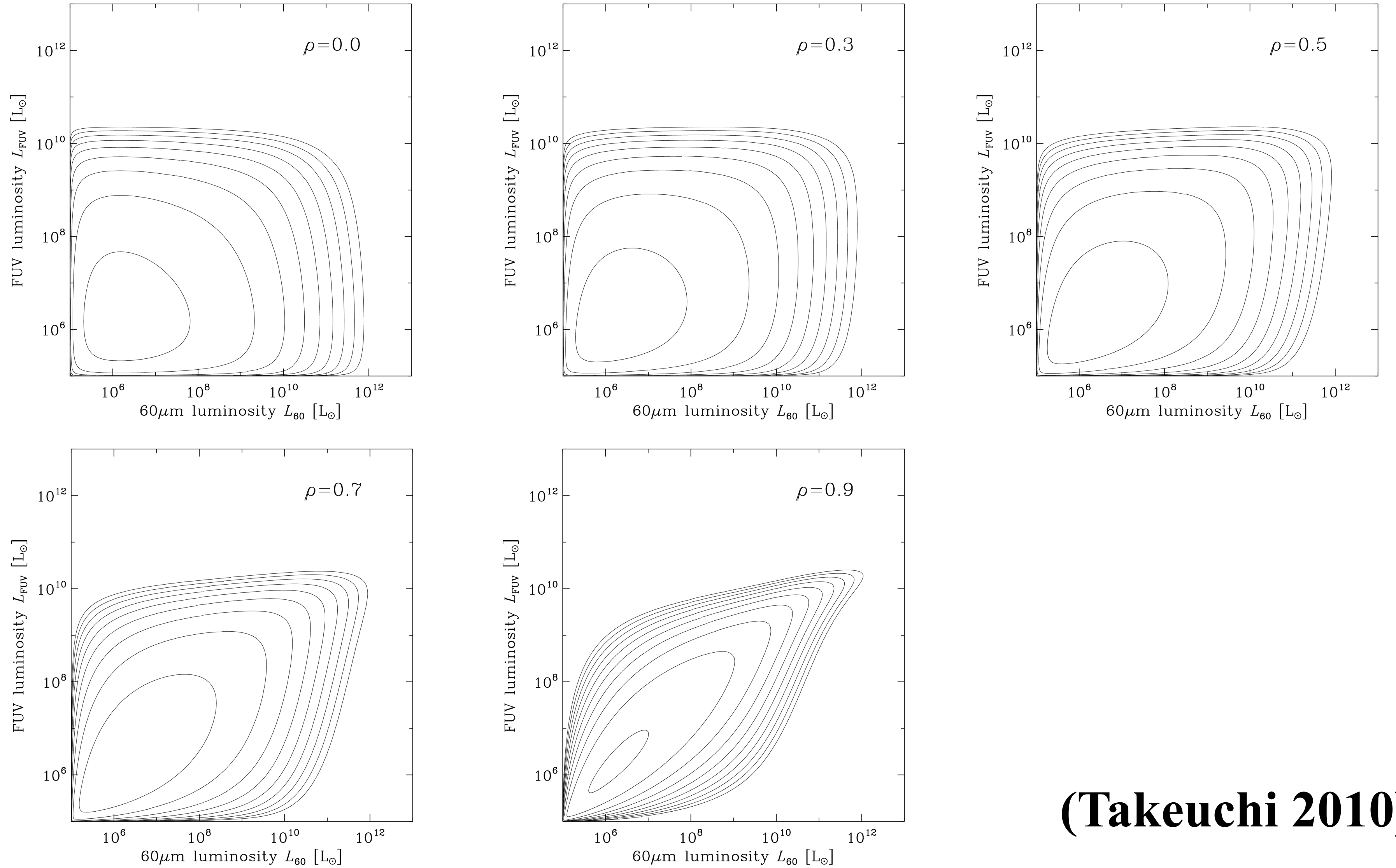
$$c^{\mathrm{G}}(u_1, u_2; \rho) \equiv \frac{\partial^2 C^{\mathrm{G}}(u_1, u_2; \rho)}{\partial u_1 \partial u_2} = \frac{\psi_2(x_1, x_2; \rho)}{\psi_1(x_1)\psi_1(x_2)} \tag{8}$$

The cumulative BLF constructed with the Gaussian copula is then expressed as

$$\Phi^{(2)}(L_1, L_2; \rho) \equiv \Psi_2\left[\Psi_1^{-1}\left(\Phi_1^{(1)}(L_1)\right), \Psi_1^{-1}\left(\Phi_2^{(1)}(L_2)\right); \rho\right] \tag{9}$$

The differential BLF is obtained by differentiating eq.(9).

The Gaussian BLFs: example



(Takeuchi 2010)

The shape of the Gaussian copula BLF depends strongly on ρ.

Benefit of copula: incorporating observational selection effects

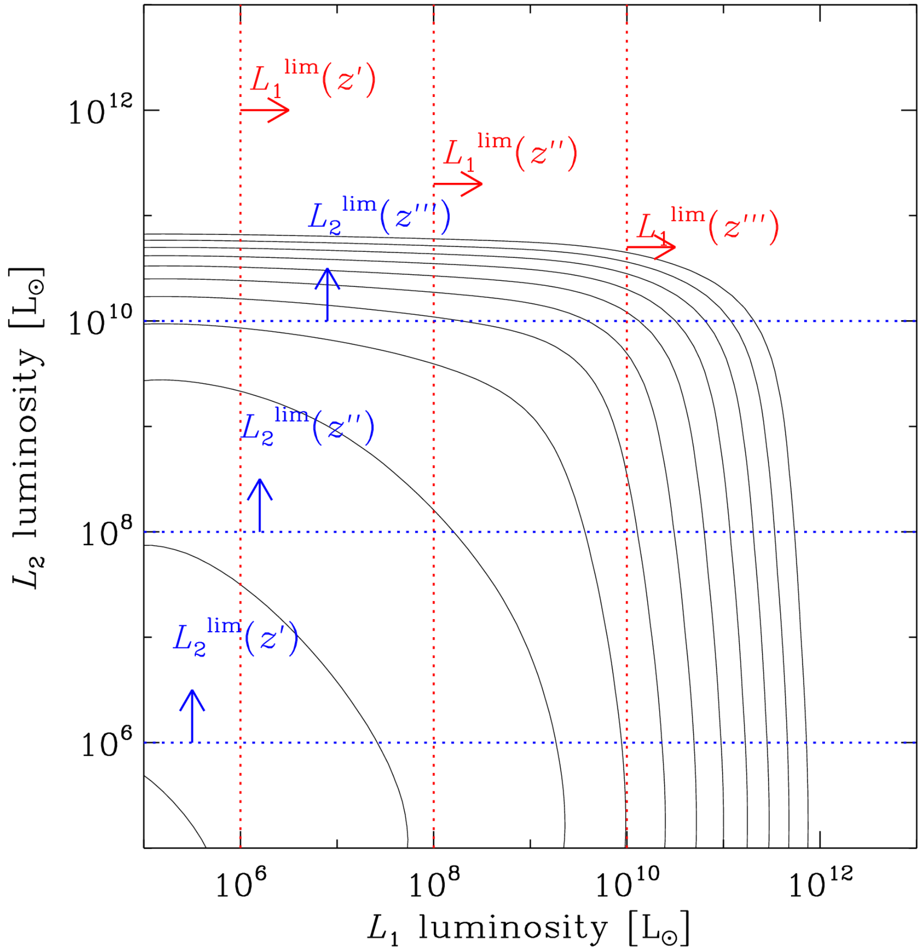
Selection effect: always exists in any kind of astronomical data.

In a bi(multi)variate analysis, there are two categories of observational selection effects.

1. **Truncation**
   We do not know if a source would exist below a detection limit.
2. **Censoring**
   We know there is a source, but we have only an upper (sometimes lower) limit for a certain observable.

We have to deal with both of these selection effects to construct a BLF from observed data at the same time. We should be careful especially when we use multiwavelength datasets.

Benefit of copula: incorporating observational selection effects



With a copula BLF, we can take into account various kind of selection effects properly (even though the formulation is messy!).

3 Bivariate Luminosity Function Analysis: Result

UV-IR bivariate LF from $z = 0$ to $z = 1$

Using the Gaussian copula, now we can estimate the bivariate luminosity function (BLF). The visible and hidden SFRs should be directly reflected to this function.

Dust is produced by SF activity, but also destroyed by SN blast waves as a result of the SF. Many physical processes are related to the evolution of the dust amount. Thus, first of all, we should *describe statistically how it evolved.*

Local samples: *IRAS, GALEX* (UV, IR-selected) + redshifts (644)
**   *AKARI, GALEX* (IR-selected) + redshifts (4086)**
**   (see A. Sakurai's poster)**

High-z samples: *Spitzer, GALEX* (UV, IR-selected) + redshifts
**   (z = 0.7, 1.0) (~ 350 for each redshift bin)**

Copula likelihood for the BLF estimation

Since we have already estimated the univariate LF at each band, we use these LFs as *given marginals.* We then estimate only one parameter, the linear correlation ρ by the likelihood below (*j*: upper limit flag, 0:detection, -1: upper limit).

$$\ln L\left(\rho \middle| L_{\mathrm{FUV},i}^{jk}, L_{\mathrm{TIR},i}^{jk}; i_k = 1,\cdots n_k\right) = \sum_{\substack{k=1:\,\mathrm{UV\text{-}sel},\\ -1:\,\mathrm{IR\text{-}sel}}} \sum_{i_k=1}^{n_k} \ln p\left(L_{\mathrm{FUV},i_k}^{jk}, L_{\mathrm{TIR},i_k}^{jk}\right)^{1+j} +$$

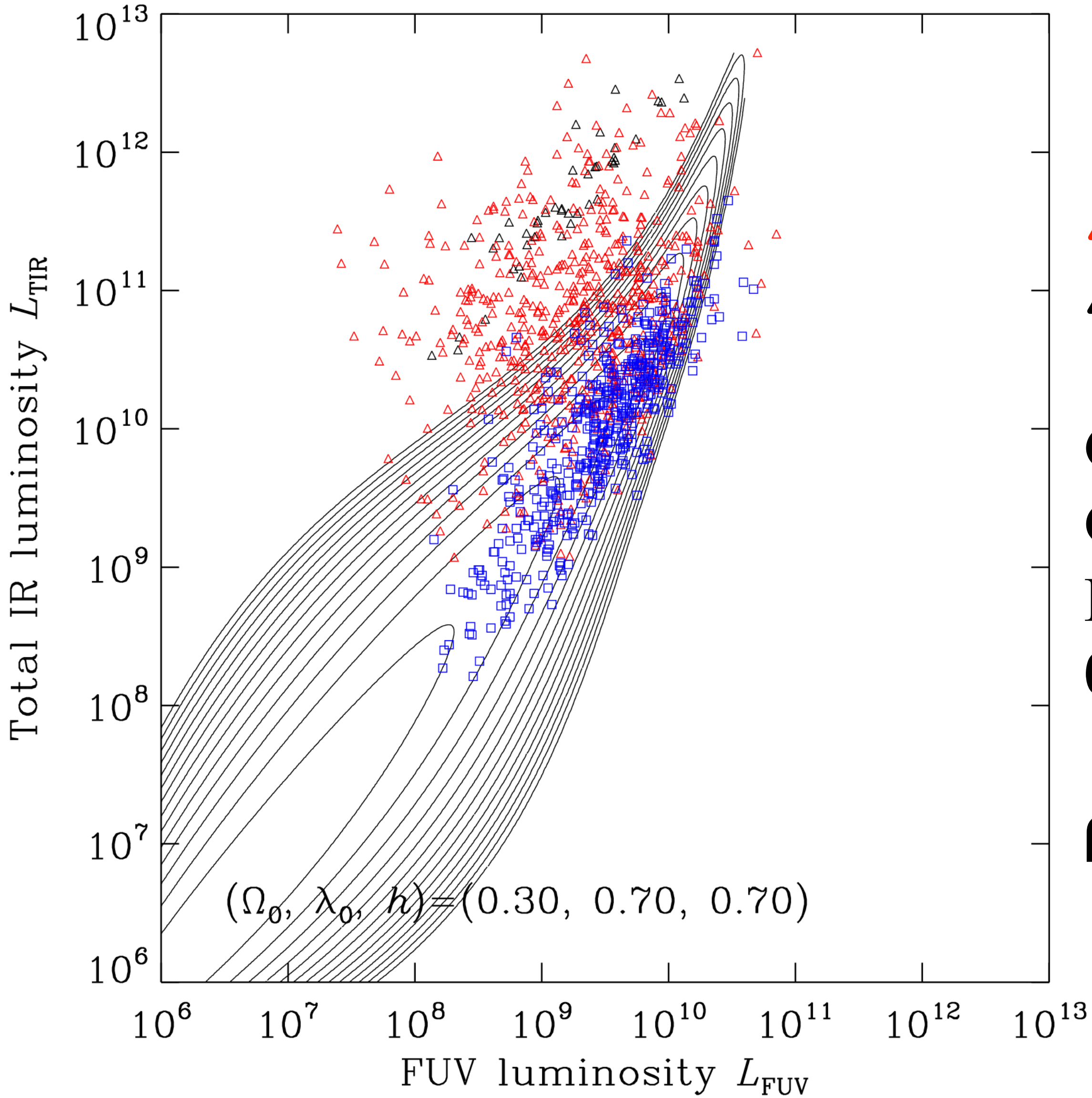
$$\left[\ln \int_0^{L_{\mathrm{FUV},i_k}^{jk}} p\left(L'_{\mathrm{FUV}}, L'_{\mathrm{TIR}}\right) dL'_{\mathrm{FUV}}\right]^{(1+k)/2} +$$

$$\left[\ln \int_0^{L_{\mathrm{TIR},i_k}^{jk}} p\left(L'_{\mathrm{FUV}}, L'_{\mathrm{TIR}}\right) dL'_{\mathrm{TIR}}\right]^{(1-k)/2}$$

These terms are necessary to treat information from upper limits.

$$p\left(L_{\mathrm{FUV},i_k}, L_{\mathrm{TIR},i_k}\right) \equiv \frac{\phi^{(2)}\left(L_{\mathrm{FUV},i_k}, L_{\mathrm{TIR},i_k}; \rho\right)}{\int_{L_{\mathrm{FUV\ or\ TIR},\lim,i_k}}^{\infty} \phi^{(2)}\left(L_{\mathrm{FUV}}, L_{\mathrm{TIR}}; \rho\right) dL_{\mathrm{FUV\ or\ TIR}}}$$

The bivariate LF at $z = 0$ (*IRAS*-*GALEX* sample)



□ : FUV-sel

□ : FUV-sel (UL at FIR)

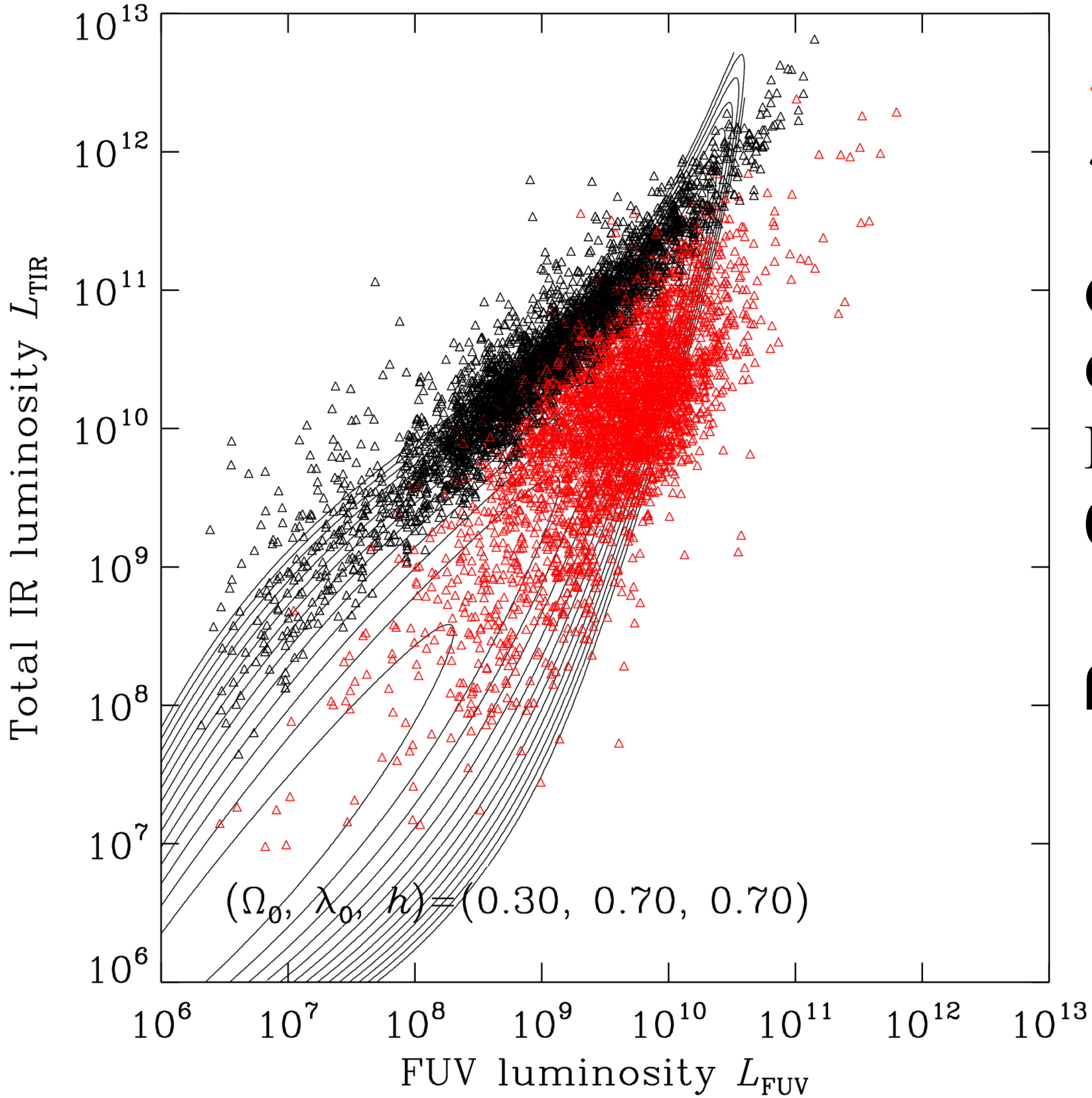
△ : FIR-sel

△ : FIR-sel (UL at FUV)

Contour:
Gaussian copula with the FUV and TIR LFs at $z = 0$.

$\rho = 0.95 \pm 0.04$

The bivariate LF at $z = 0$ (*AKARI*-*GALEX* sample)



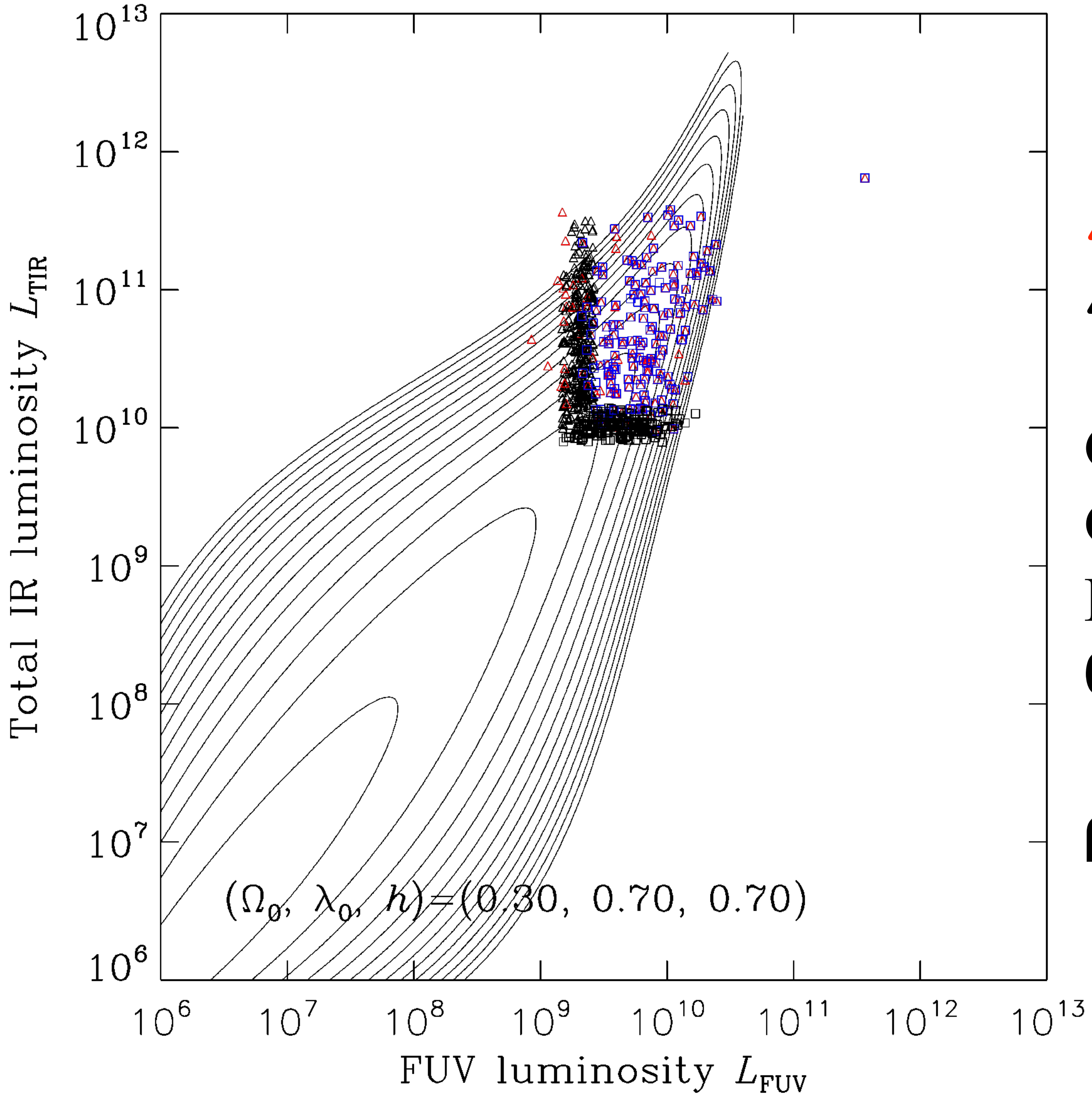
△ : FIR-sel

△: FIR-sel (UL at FUV)

Contour:
Gaussian copula with the FUV and TIR LFs at $z = 0$.

$\rho = 0.95 \pm 0.006$

The bivariate LF at z = 0.7 (*Spitzer-GALEX* sample)



□ : FUV-sel

□ : FUV-sel (UL at FIR)

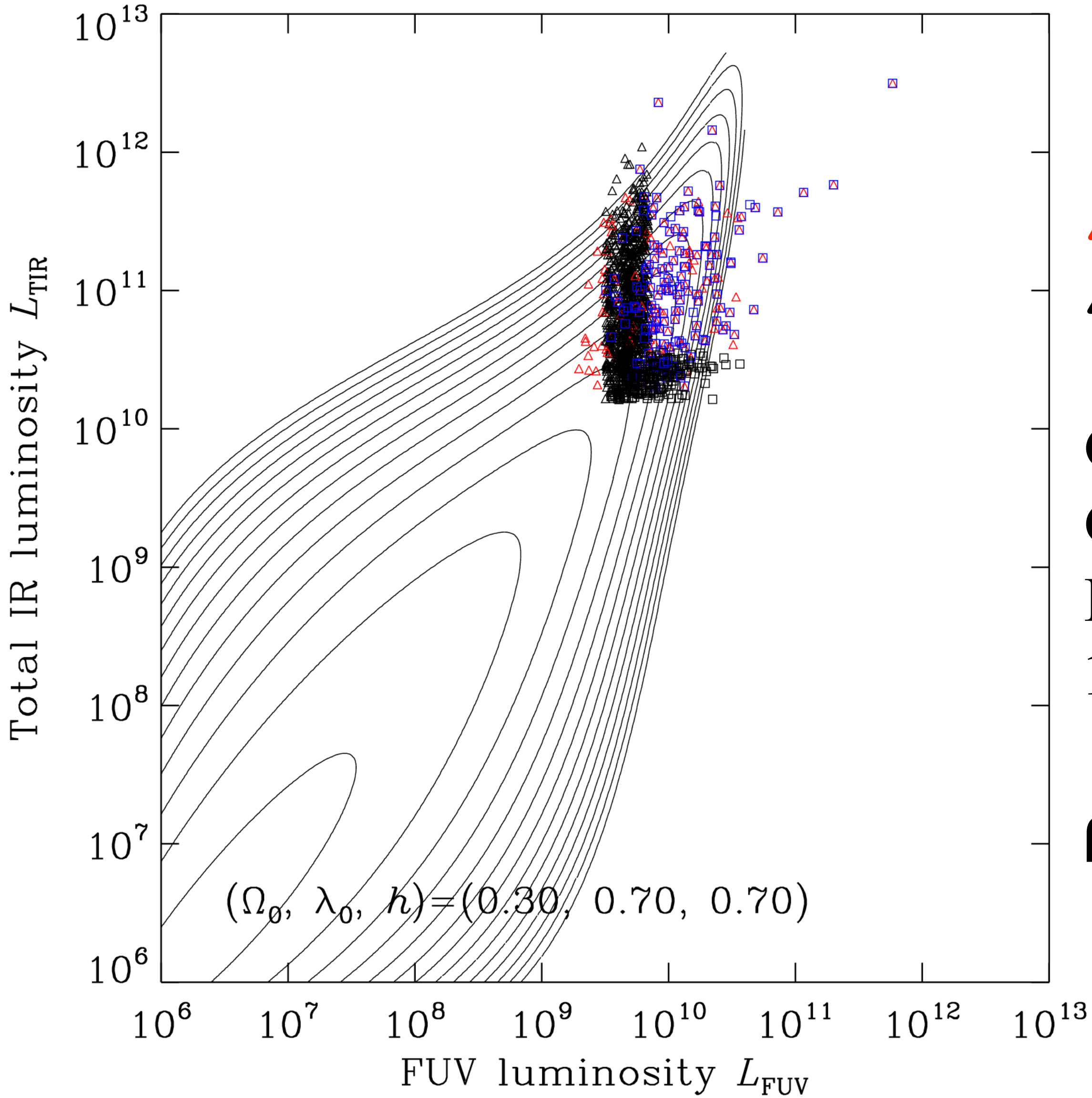
△ : FIR-sel

△ : FIR-sel (UL at FUV)

Contour:
Gaussian copula with the FUV and TIR LFs at z = 0.7.

ρ = 0.91±0.05

The bivariate LF at z = 1.0 (*Spitzer*-*GALEX* sample)



□ : FUV-sel

□ : FUV-sel (UL at FIR)

△ : FIR-sel

△ : FIR-sel (UL at FUV)

Contour:
Gaussian copula with the FUV and TIR LFs at z = 1.0.

ρ = 0.85±0.05

Result from the copula BLF analysis

In the Local Universe, the BLF is quite well constrained. It is rather impressive that the estimated correlation coefficient ρ is very high ~ 0.95, both from *IRAS*-*GALEX* and *AKARI*-*GALEX* datasets.
The apparent scatter of the L_{FUV}-L_{TIR} is found to be due to the nonlinear shape of the ridge of the BLF. This bent shape of the BLF was implied by preceding studies (e.g., Martin et al. 2005). The copula BLF naturally reproduced this.

At higher redshifts (z = 0.7-1.0), the linear correlation remains tight (ρ ~ 0.85-0.9) even though it is difficult to constrain the low-luminosity end from the data in this analysis (*Spitzer*-*GALEX* in the CDFS). It will be interesting to apply this method to better forthcoming data.

4 Summary

To understand the visible and hidden star formation history in the Universe, it is crucial to analyze multiwavelength data in a coherent and synthesized manner.

1. **The copula method is an ideal tool to combine two (or more) marginal univariate LFs to construct a bi(multi-)variate LFs.**
2. **Copula is also useful to incorporate selection effects.**
3. **The Gaussian copula LF is sensitive to the linear correlation parameter ρ.**
4. **Even so, ρ in the copula LF is remarkably stable with redshifts (from 0.95 at $z = 0$ to 0.85 at $z = 1.0$).**
5. **This implies the evolution of the UV-IR bivariate LF is mainly due to the different evolution of the univariate LFs, and may not be controlled by the dependence structure.**

The data used in this work are not deep enough, but *Herschel* and ALMA data will improve the estimates drastically.